ISSN 1561-5359. Штучний інтелект, 2016, № 3

УДК 004.92.93.94

С.С. Кондратюк, Ю.В. Крак
Київський національний університет імені Тараса Шевченка, Україна
вул. Володимирська, 64/13, м. Київ, 01601

ПЛАТФОРМОНЕЗАЛЕЖНЕ ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ДЛЯ РОЗРОБКИ СИСТЕМ ЖЕСТОВОЇ КОМУНІКАЦІЇ: МОДЕЛЮВАННЯ ДАКТИЛЬНОЇ МОВИ

S.S. Kondratiuk, Iu.V. Krak
Taras Shevchenko National University of Kyiv, Ukraine
Volodymyrska str., 64/13, Kyiv, 01601

CROSS PLATFORM SOFTWARE FOR THE DEVELOPMENT OF SIGN COMMUNICATION SYSTEM: DACTYL LANGUAGE MODELLING

Запропоновано технологію, що реалізується за допомогою кросплатформених засобів, для моделювання жестових одиниць жестової мови, динамічного відображення станів між жестовими одиницями при поєднанні їх у жестові конструкції (слова, речення). Технологія реалізує відтворення змодельованих жестових одиниць і конструкцій за допомогою просторової віртуальної моделі руки. За допомогою кросплатформених засобів, технологія має можливість запуску на множині платформ без імплементації під кожну платформу.
Ключові слова: кросплатформеність, жестова мова, моделювання жестів.
The technology is proposed, which is implemented using cross-platform tools, for modeling of signs units of sign language, dynamic mapping between states of sign units with a combination of gestures structures (words, sentences). The technology realizes simulated playback of signed items and constructions using virtual model of spatial hand. With the cross-platform means technology achieves the ability to run on multiple platforms without implementation for each platform.
Key words: cross-platform, sign language, sing modeling.

Вступ

Спілкування за допомогою жестів є одним з трьох основних засобів передачі інформації між людьми, доповнюючи символьну (текстову) та голосову (мовленнєву) комунікації. Жестову мову зазвичай використовують люди з вадами слуху для спілкування між собою, та зі своїм оточенням, що збільшує коло людей яким необхідно володіти нею. Відзначимо, що жестова мова є універсальною в тому розумінні, що текстову інформацію можна передати жестами, а коли жестів для певних слів не існує (імена, назви міст, територій тощо), то їх показ здійснюється по буквах за допомогою дактильної абетки.

Сучасні технічні засоби дозволяють швидко та майже без обмежень накопичувати інформацію, оброблювати її як у хмарних обчисленнях (модель, що забезпечує повсюдний, зручний доступ за запитом через мережу до віртуального кластеру обчислювальних ресурсів) [1], так і локально на пристрої, та завдяки швидким каналам передачі даних повертати результати обчислень до користувача. Все це актуально й для жестової мови, для накопичення та відтворення якої тепер доступний широкий загал різноманітних пристроїв та платформ, як стаціонарних так і мобільних, як потужних, так і енергоефективних. Актуальною проблемою є перенесення відтворення жестової мови на всі ці пристрої, для використання людьми з вадами слуху зокрема та всіма користувачами в цілому. У даній проблемі головним є розгортання уніфікованої технології на різноманітті доступних платформ (android, ios, windows, linux, web) без необхідності портування та переписування з початку.

Одним із засобів вирішення поставленої проблеми з візуалізації та відтворення жестової мови є кросплатформена розробка та кросплатформене програмне забезпечення. На відміну від одноплатформенних технологій, які функціонують лише на специфічній платформі під яку вони розроблялись, "кросплатформеність надає програмному забезпеченню можливість виконання на більш ніж одній платформі з однаковою (або майже однаковою) функціональністю" [2]. Термін "платформа" в даному контексті може відноситися до одного або комбінації кількох визначень: 1) тип операційної системи (наприклад Microsoft Windows, Mac OS X, Linux, Solaris, Android, iOS); 2) тип процесора (наприклад x86, PowerPC, ARM); 3) тип апаратного забезпечення (наприклад, mainframe, робоча станція, персональний комп'ютер, мобільний пристрій) [2]. Кросплатформені технології стоять в одному ряду із технологіями незалежними від платформи (такі, що здатні функціонувати на будь-якій платформі, наприклад веб-додатки) [2] та кросплатформеними віртуальними машинами (технології, що підтримують індивідуальні процеси чи системи залежно від рівня абстракції, на якому відбувається віртуалізація) [3].

У даній статті запропоновано вирішення проблеми за допомогою кросплатформеності, з урахуванням характеристик різних класів пристроїв (типу апаратного забезпечення, потужності центрального процесора, кількості оперативної пам'яті, наявності мережі Інтернет) та налаштування рівня деталізації тривимірної моделі кисті та пальців руки людини з анімацією їх рухів.

Таким чином, кросплатформена інформаційно-комунікаційна технологія для моделювання дактильної жестової мови є актуальною та важливою проблемою у плані розробки нових підходів до побудови кросплатформених систем та їх алгоритмічної та програмної реалізації.

Існуючі підходи до моделювання жестової мови та їх реалізації на різних платформах

Моделювання жестів є проблемою, яка розглядалася як самостійно, так і у рамках задачі моделювання і розпізнавання жестів та побудови комп'ютерних засобів для вивчення жестової мови. Однією з систем для відображення жестової мови є American Sing Language Online Dictionary [4], яка складається з бази відеозаписів слів та фраз показаних жестовою мовою. Відзначимо також, що керування за допомогою жестів є актуальною проблемою у розробках платформонезалежних систем людино-машинного інтерфейсу [5]. Такими розробками займається ряд комерційних установ (див., наприклад, [6][7]), проте запропоновані ними системи налаштовані на ряд заздалегідь визначених непов'язаних жестів, а, отже, не розв'язують проблему моделювання жестової мови.

Створення моделі руки є першим кроком у задачі моделювання жестів. У статті [8] проводиться аналіз існуючих підходів подання моделі руки, які поділяються на дві основні групи: просторові та часові. Перші розглядають характеристики позицій руки для різних жестів, другі відносяться до опису динаміки жестів. Моделювання руки у просторовій області може бути виконане у дво- та тривимірному просторі.

У статті [9] наведено систему, здатну моделювати жестові анімації для заданого тексту. В рамках даної системи було використано статистичну модель для аналізу вхідного тексту та генеративний алгоритм для створення відповідних кінематично симульованих жестових анімацій. У рамках статті, автори надали інструмент ANVIL для анотації вхідного тексту, генератор жестів NOVA, та для анімації жестів використовують бібліотеку DANCE, розроблену в роботі [10]. Як і використана бібліотека, так і розробка авторів, побудовані на платформі Microsoft Windows і x86 процесора. У статті [11] автори розглядають моделювання рухів віртуального

персонажа для просторового відтворення жестової мови на платформі Microsoft Windows. У статті [12], розроблена система навчання жестовій мові, яка складається з двох модулів - модуля демонстрації жестів за допомогою відеозаписів та модуля розпізнавання жесту продемонстрованого користувачем (необхідні рукавиці), в основі якого використано приховану марковську модель. Система навчання побудована на платформі Microsoft Windows і x86 процесора.

Розпізнавання жестів на мобільних платформах розробляється у статті [13], але моделювання жестів на мобільних пристроях не проводиться.

Постановка задачі

Необхідно запропонувати технологію для моделювання жестових одиниць (морфем, візем, хірем [14]) жестової мови, динамічного відображення станів між жестовими одиницями при поєднанні їх у жестові конструкції (слова, речення) та реалізувати відтворення змодельованих жестових одиниць і конструкцій за допомогою просторової віртуальної моделі руки.

Технологія повинна розв’язати проблему запуску на вже існуючій множині платформ за допомогою кросплатформеності, без імплементації функціоналу під кожну платформу окремо. Ефективність запропонованого підходу буде показана на побудові кросплатформеної технології для моделювання та розпізнавання елементів дактильної абетки української жестової мови.

Запропонований підхід

Для вирішення проблеми моделювання жестових одиниць жестової мови та динамічного відображення жестових конструкцій за допомогою просторової віртуальної моделі руки запропоновано кросплатформену технологію яка базується на кросплатформеній біблотеці libHand [15] (для моделювання руки та кросплатформеній бібліотеці Qt [16] для користувацького інтерфейсу, обидві бібліотеки та технологію імплементовано мовою програмування C++ (Рис. 1). Запропоновані засоби дозволяють вирішити проблему запуску технології на існуючій множині платформ. Новизна запропонованої технології полягає у кросплатформеності та можливості налаштування рівня деталізації тривимірної моделі кисті та пальців руки людини та анімації їх рухів, що надає можливість запуску технології без змін на множині платформ (різних типах процесорів, операційних систем та апаратному забезпеченні).

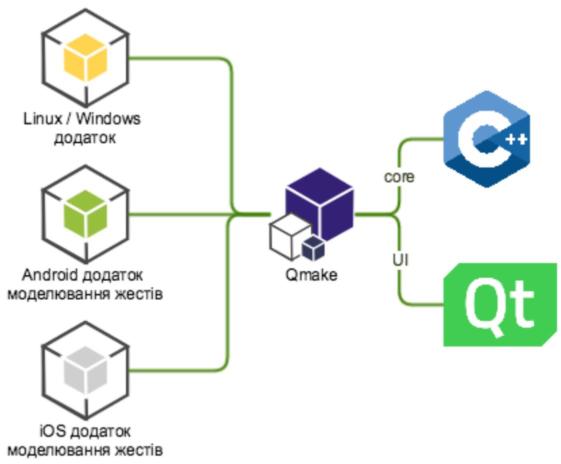


Рис. 1. Кросплатформеність та використані технології/бібліотеки

Перевагою кросплатформених технологій над технологіями що розробляються для єдиної платформи є відсутність необхідності модифікувати чи імплементувати заново вже наявний функціонал для іншої платформи (портування) [2], що пришвидшує процес розробки та розгортання технології, та збільшує коло потенційних користувачів.

Перевагою кросплатформених технологій над кросплатформеною віртуальною машиною та емуляторами, зокрема, є швидкодія та відсутність необхідності у встановленні додаткового програмного забезпечення (програмних залежностей).

Моделювання руки

Для побудови моделі руки людини скористаємось тим, що рука має 27 кісток, з них 8 кісток знаходяться в зап'ясті, 3 знаходяться у великому пальці (одна п'ястковa та 2 фаланги), і 4 п'ясткові та 12 фаланг знаходяться у інших пальцях [17]. Кожна кістка з'єднується з іншими за допомогою різних видів суглобів.

Розробка власного кросплатформеного движку для моделювання руки є нетривіальною задачею, тому у якості основної технології для моделювання тривимірної моделі руки та анімацій переходу між морфемами було обрано кросплатформену бібліотеку libHand. Бібліотека libHand здатна ефективно відтворювати та деформувати реалістичну модель руки яка складається з понад 70000 полігонів. Перевагою libHand також є інтерфейс для бібліотеки OpenCV, що надає можливість виконувати обробку моделі руки засобами computer vision.

Базуючись на анатомії руки, у рамках libHand було розроблено модель руки з 25 ступенями рухливості, чотири з них розташовані у п'ястно-зап'ястному суглобі, для мізинця та перстеневого пальця, щоб забезпечити дуговидний рух долоні. Великий палець має 5 ступенів рухливості, середній та вказівний пальці мають по 4 ступені рухливості (п'ястно-фаланговий суглоб має два ступеня рухливості, та дистальний і проксимальній міжфалангові суглоби мають кожний по одному).

Для збереження жестів було обрано формат YAML [18], в якому жест зберігається в наступному вигляді:

%YAML:1.0 rotation: [-9.5845758914947510e-02,
8.3791027449819921e-09, -9.9539619684219360e-01,
-7.5690525770187378e-01, -6.4944797754287720e-01,
7.2881683707237244e-02, -6.4645802974700928e-01,
7.6040601730346680e-01, 6.2246840447187424e-02]
hand_joints: finger1joint1: [0., 0., 0.] finger1joint2: [0., 0., 0.]...

Моделювання жестів

Завдяки даній технології реалізовано "просунутий" модуль для модифікації існуючих жестів жестової мови та додавання нових жестів жестових мов інших абеток, що надає потенціал для локалізації та безперешкодного використання інформаційної технології для абеток інших мов (до та після модифікації, рис. 2).

Даний модуль модифікації та додавання жестів є доступним для апаратного забезпечення з встановленими наступними операційними системами: macOS, Linux, Windows.

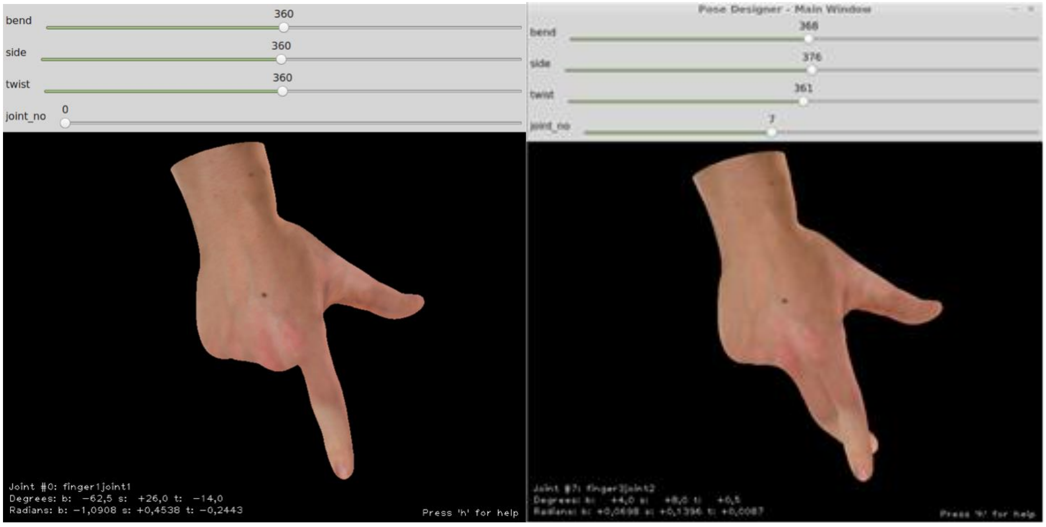


Рис. 2. Моделювання жестів

Розробка інтерфейсу користувача

З міркувань швидкодії та кросплатформеності, технологією для користувацького інтерфейсу було обрано бібліотеку Qt. Бібліотека Qt надає можливість розробити масштабований та кросплатформений інтерфейс, перевагою якого над існуючими рішеннями є швидкодія та уніфікований вигляд на множині існуючих платформ, як мобільних так і стільникових. Користувацький інтерфейс для технології жестової комунікації реалізовано за допомогою Qt (рис. 3).



Рис. 3. Інтерфейс користувача

Використання кросплатформених засобів

Запропонована інформаційна технологія, завдяки обраним кросплатформеним засобам для імплементації, розв'язує проблему запуску на існуючій множині платформ без реалізації під кожну платформу окремо.

Програмні засоби запропоновані та використані під час імплементації інформаційної технології є кросплатформеними та функціонують без змін незалежно від операційної системи (Windows, Linux, Android, iOS), типу процесора (x86, arm), та типу апаратного забезпечення (мобільний чи стаціонарний пристрій).

Завдяки кросплатформеній системі побудови додатку QMake [19] можливо створити додатки для кожної платформи без необхідності портування чи зміни оригінального коду (рис. 1). Оскільки специфічних вимог до апаратного забезпечення у інформаційної технології для моделювання жестової мови немає, то існують об'єктивні перешкоди у швидкодії пристроїв старших поколінь. Для подолання даної проблеми запропоновано наступний адаптивний підхід у інформаційній технології (рис. 4).

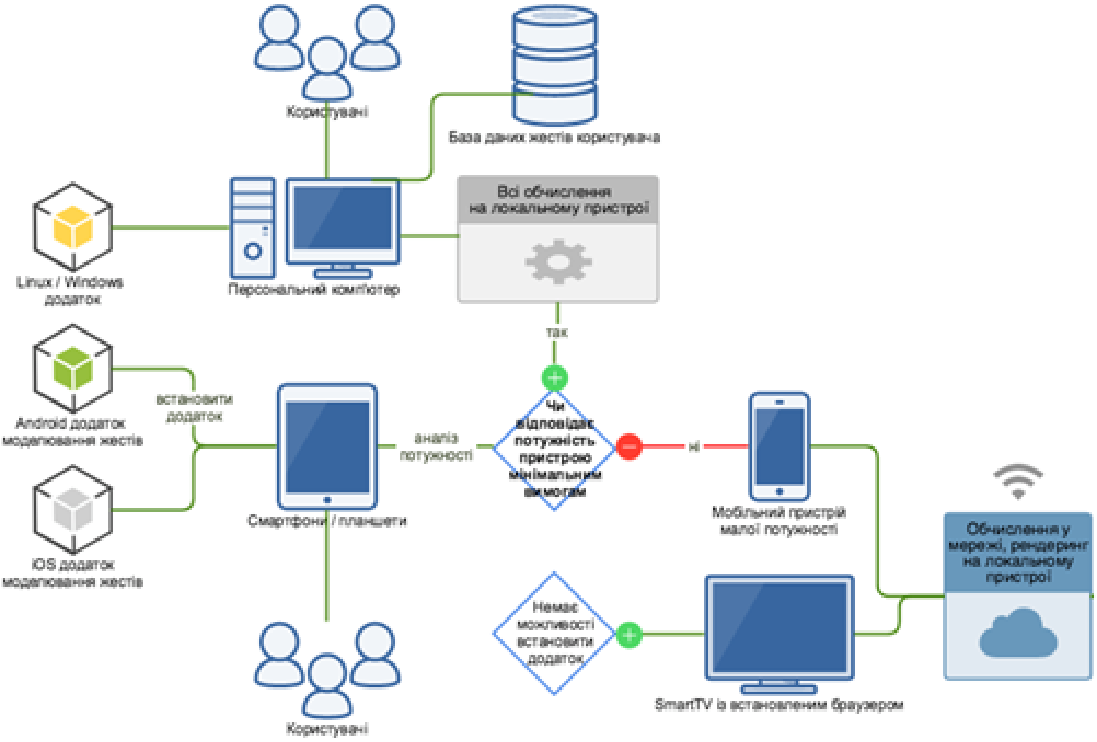


Рис. 4. Загальна схема кросплаторменості та адаптивного запуску інформаційної технології

У випадку мобільного додатку (iOS, Android) чи додатку на пристрою зі стаціонарною ОС (Windows, Linux), при інсталяції на пристрій, інформаційна технологія проводить аналіз наявного апаратного забезпечення та, залежно від його потужності, проводить ряд налагоджень: 1) полігональність моделі руки змінюється, з пріоритетом до швидкодії; 2) під час повороту моделі руки змінюється крок кута, на який вона обертається, з пріоритетом до швидкодії (рис. 5). У випадку, якщо наявне апаратне забезпечення не відповідає мінімальним вимогам інформаційної технології, користувачу надається рекомендація вибрати "онлайновий" режим роботи, в якому обчислення не виконуються на апаратному забезпеченні.

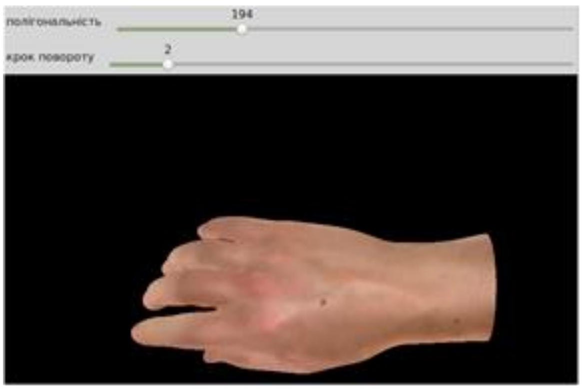


Рис. 5. Зміна полігональності моделі руки

Висновки

Запропоновано технологію, побудовану кросплатформеними засобами, для моделювання жестових одиниць жестової мови, динамічного відображення станів між жестовими одиницями при поєднанні їх у жестові конструкції. Технологія реалізує відтворення змодельованих жестових одиниць і конструкцій за допомогою просторової віртуальної моделі руки.

Технологія за допомогою кросплатформеності розв'язує проблему запуску на існуючій множині платформ, без створення функціоналу під кожну платформу окремо. Таким чином, було показано ефективність технології, побудованої за допомогою кросплатформених засобів, на прикладі моделювання та розпізнавання елементів дактильно азбуки української мови жестів.

Інформаційно-комунікаційна технологія розроблялася з можливостями подальшого масштабування як на рівні жестів однієї абетки, так на рівні жестів інших абеток. Для реалізації даної ідеї може бути використано механізм валідації нових жестів до загальної бази даних жестів. Кросплатформеність інформаційно-комунікаційної технології та уніфікований протокол та формат передачі даних (YAML) дозволяє використовувати ряд рішень для віддалених обчислень за допомогою хмарних обчислень, веб-серверів, локальних серверів використовуючи єдину базу даних жестів PostgreSQL [20].

Література

1. Mell P., Grance T. The NIST Definition of Cloud Computing (Technical report)/National Institute of Standards and Technology: U.S. Department of Commerce. doi:10.6028/NIST.SP.800-145.
2. The Linux Information Project, Cross-platform Definition [Електр.ресурс]. - Режим доступу: http://www.linfo.org/cross-platform.html
3. Smith J., Nair R. The Architecture of Virtual Machines // Computer. IEEE Computer Society. – Vol. 38. – No5. – 2005. – P. 32–38.
4. ASL Sing language dictionary [Електр.ресурс]. - Режим доступу: http://www.signasl.org/sign/model
5. Гращенко Л.А., Фисун А.П. и др. Теоретические и практические основы человеко-компьютерного взаимодействия: базовые понятия человеко-компьютерных систем в информатике и информационной безопасности: Монография / Ред. А.П. Фисун. – Орел: ОГУ, 2004. –169 с.
6. Samsung TV Gesture book [Електр.ресурс]. - Режим доступу: http://www.samsung.com/ph/smarttv/common/guide_book_3p_si/waving.html
7. Apple Touchless Gesture System for iDevices [Електр.ресурс]. - Режим доступу: http://www.patentlyapple.com/patently-apple/2014/12/apple-invents-a-highly-advanced-air-gesturing-system-for-future-idevices-and-beyond.html
8. Khan R.Z., Ibraheem N.A., Meghanathan N. et al. (Eds). Comparative study of hand gesture recognition system (SIPM, FCST, ITCA, WSE, ACSIT, CS & IT 06, 2012. – P. 203–213.
9. Neff M., Michael M., Albrecht I., Seidel H.-P. Gesture Modeling and Animation by Imitation // MPI–I–2006–4-008 September 2006
10. Shapiro A., Chu D., Allen B., Faloutsos P. Dynamic Controller Toolkit [Електр.ресурс]. - Режим доступу: http://www.arishapiro.com/Sandbox07_DynamicToolkit.pdf
11. Кривонос Ю.Г. Моделювання рухів віртуального персонажа для просторового відтворення жестової мови. / Ю.Г. Кривонос, Ю.В. Крак, О.В. Бармак // Реєстрація, зберігання і обробка даних. – Т. 12, № 2. – 2010. – С. 181-189.
12. Aran O. Sign Language Tutoral Tool // eNTERFACE'06. July 17 – August 11, 2006. Final Project Report. – Dubrovnic, Crotia.
13. Jagdish L. Raheja. Android based portable hand sign recognition system // March 2015 DOI: 10.15579/gcsr.vol3.ch1 · Source: arXiv
14. Stokoe W. C. 1960. Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf . 1969. – P.61-67.
15. LibHand library [Електр.ресурс]. - Режим доступу: https://github.com/jonkeane/libhand
16. Qt library documentation [Електр.ресурс]. - Режим доступу: https://www.qt.io/developers/

17. Tubiana R, Thomine J, Mackin E. Examination of the handand wrist. 2nd ed. Martin Dunitz. ISBN: 1853175447/1-85317-544-7. Publisher: Informa Healthcare. 1996.
18. YAML – The Official YAML Web Site [Електр.ресурс]. - Режим доступу: http://yaml.org/
19. Qmake build tool manual [Електр.ресурс]. - Режим доступу: http://doc.qt.io/qt-4.8/qmake-manual.html
20. PostgreSQL official web site [Електр.ресурс]. - Режим доступу: https://www.postgresql.org/

Literatura

1. Mell P., Grance T. The NIST Definition of Cloud Computing (Technical report)/National Institute of Standards and Technology: U.S. Department of Commerce. doi:10.6028/NIST.SP.800-145.
2. The Linux Information Project, Cross-platform Definition [Електр.ресурс]. - Режим доступу: http://www.linfo.org/cross-platform.html
3. Smith J., Nair R. The Architecture of Virtual Machines // Computer. IEEE Computer Society. – Vol. 38. – No5. – 2005. – P. 32–38.
4. ASL Sing language dictionary [Електр.ресурс]. - Режим доступу: http://www.signasl.org/sign/model
5. Graschenko L.A., Fisun A.P. i dr. Teoreticheskie i prakticheskie osnovyi cheloveko-kompyuternogo vzaimodeystviya: bazovyie ponyatiya cheloveko-kompyuternyih sistem v informatike i informatsionnoy bezopasnosti: Monografiya / Red. A.P. Fisun. – Orel: OGU, 2004. –169 s.
6. Samsung TV Gesture book [Електр.ресурс]. - Режим доступу: http://www.samsung.com/ph/smarttv/common/guide_book_3p_si/waving.html
7. Apple Touchless Gesture System for iDevices [Електр.ресурс]. - Режим доступу: http://www.patentlyapple.com/patently-apple/2014/12/apple-invents-a-highly-advanced-air-gesturing-system-for-future-idevices-and-beyond.html
8. Khan R.Z., Ibraheem N.A., Meghanathan N. et al. (Eds). Comparative study of hand gesture recognition system (SIPM, FCST, ITCA, WSE, ACSIT, CS & IT 06, 2012. – P. 203–213.
9. Neff M., Michael M., Albrecht I., Seidel H.-P. Gesture Modeling and Animation by Imitation // MPI–I–2006–4-008 September 2006
10. Shapiro A., Chu D., Allen B., Faloutsos P. Dynamic Controller Toolkit [Електр.ресурс]. - Режим доступу: http://www.arishapiro.com/Sandbox07_DynamicToolkit.pdf
11. Kryvonos Yu.H. Modelyuvannya rukhiv virtual'noho personazha dlya prostorovoho vidtvorennya zhestovoyi movy. / Yu.H. Kryvonos, Yu.V. Krak, O.V. Barmak // Reyestratsiya, zberihannya i obrobka danykh. – T. 12, No 2. – 2010. – C. 181-189.
12. Aran O. Sign Language Tutoral Tool // eNTERFACE’06. July 17 – August 11, 2006. Final Project Report. – Dubrovnic, Croatia.
13. Jagdish L. Raheja. Android based portable hand sign recognition system // March 2015 DOI: 10.15579/gcsr.vol3.ch1 · Source: arXiv
14. Stokoe W. C. 1960. Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf . 1969. – P.61-67.
15. LibHand library [Електр.ресурс]. - Режим доступу: https://github.com/jonkeane/libhand
16. Qt library documentation [Електр.ресурс]. - Режим доступу: https://www.qt.io/developers/
17. Tubiana R, Thomine J, Mackin E. Examination of the handand wrist. 2nd ed. Martin Dunitz. ISBN: 1853175447/1-85317-544-7. Publisher: Informa Healthcare. 1996.
18. YAML – The Official YAML Web Site [Електр.ресурс]. - Режим доступу: http://yaml.org/
19. Qmake build tool manual [Електр.ресурс]. - Режим доступу: http://doc.qt.io/qt-4.8/qmake-manual.html
20. PostgreSQL official web site [Електр.ресурс]. - Режим доступу: https://www.postgresql.org/

RESUME

S.S. Kondratiuk, Iu.V. Krak

Cross platform software for the development of sign communication system: dactyl language modelling

Communicating with gestures is one of the three main means of information transmission between people, as well as a character (text) and voice communication. Sign language is universal in the sense that the text information can be transmitted by gestures, as well as if gestures for certain words do not exist (names, names of cities, areas, etc.), their demonstration is performed by using the letters of the dactyl alphabet.

Modern technical equipment allows to quickly and almost without any restrictions to accumulate information and process it remotely or locally on the device, and thanks to the fast

data transfer channels to return the calculation results to the user. All this is true for sign language for the accumulation and reproduction of which it is now available a variety of devices and platforms. The authors proposed information technology to solve the problem of modeling of sign units of sign language and the dynamic display of sign structures with spatial virtual hand model. An important step in that the proposed information technology is cross-platform, so the ability to run on multiple platforms is achieved. This system is based on a cross-platform library for modeling hands libHand and cross-platform Qt library for the user interface.

The novelty of the proposed technology is its cross-platform and the possibility of setting the level of detail of the three-dimensional model of the human hand and the animation of their movements.

The existing cross-platform software tools for modeling of three-dimensional models has been researched and as the core technology to simulate a three-dimensional model of hand and transition animations between morphemes cross platform libHand library has been chosen. As a part of the technology it has been developed hand model with 25 degrees of freedom. In the proposed sign language simulation technology number of polygons in the hand model is a tunable parameter, so the technology is able to run even on low-end hardware.

As a part of the technology the module was proposed, which provides means to modify existing gestures of sign language and adding new gestures of alphabets of other languages, and as a demonstration, gestures of Ukrainian dactyl alphabet were added.

With the aim of speed and cross-platform a Qt library was proposed as a UI technology. It provides a possibility to develop a scalable and cross-platform interface, which advantages over existing solutions are performance and a unified view on the existing platforms.

If there is no hardware to meet the minimum requirements of information technology, there is provided "online" mode in which the calculations and display of gesture animation are not executed on the hardware but are performed via the network on the remote server or on the cloud computing resources.

Thus, it shown the effectiveness of the technology built using cross-platform tools for modeling and recognition of elements of Ukrainian dactyl sign language alphabet.

Надійшла до редакції 01.10.2016